

# The normal distribution

- 11A z-scores
- 11B Comparison of scores
- 11C Distribution of scores



## Syllabus reference

Data analysis 6

- The normal distribution

This chapter will focus on using measures of central tendency and spread to compare scores from different data sets and to recognise when a score within a data set is unusual.

# ARE YOU READY?

Try the questions below. If you have difficulty with any of them, extra help can be obtained by completing the matching SkillSHEET. Either click on the SkillSHEET icon next to the question on the *Maths Quest HSC Course* eBookPLUS or ask your teacher for a copy.

## eBookplus

**Digital doc**  
SkillSHEET 11.1  
doc-1408  
**Finding the mean**

### Finding the mean

- 1 Find the mean of each of the following data sets.

**a** 4, 6, 2, 7, 9, 3, 6, 9

**b** 45, 72, 90, 70, 69, 48, 62, 99, 47, 55

**c**

Score	Frequency
6	3
7	13
8	16
9	12
10	6

**d**

Score	Frequency
15	12
16	36
17	42
18	15
19	11
20	9

## eBookplus

**Digital doc**  
SkillSHEET 11.2  
doc-1409  
**Finding the standard deviation**

### Finding the standard deviation

- 2 Find the population standard deviation of each of the data sets in question 1. Give each answer correct to 1 decimal place.

## eBookplus

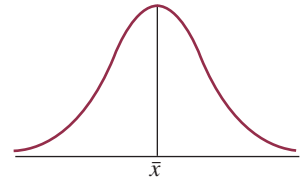
**Digital doc**  
SkillSHEET 11.3  
doc-1410  
**Choosing the appropriate standard deviation**

### Choosing the appropriate standard deviation

- 3 In each of the following choose the appropriate measure of standard deviation.
- a** On entering a certain music store people are asked how many CDs they own.
  - b** The number of parliamentarians who vote in favour of bills brought before parliament.

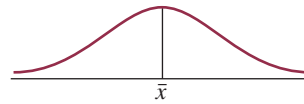
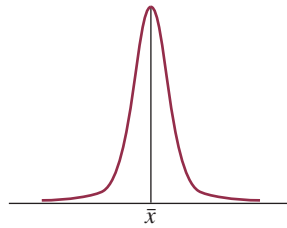
## 11A z-scores

A **normal distribution** is a statistical occurrence where a data set of scores is symmetrically distributed about the mean. Most continuous variables in a population, such as height, mass and time, are normally distributed. In a normal distribution, the frequency histogram is symmetrical and begins to take on a bell shape as shown by the figure on the right.



The normal distribution is symmetrical about the mean, which has the same value as the median and mode in this distribution. The graph of a normal distribution will extend symmetrically in both directions and will always remain above the  $x$ -axis.

The spread of the normal distribution will depend on the standard deviation. The lower the standard deviation, the more clustered the scores will be around the mean. The figure below left shows a normal distribution with a low standard deviation, while the figure below right shows a normal distribution with a much greater standard deviation.



To gain a comparison between a particular score and the rest of the population, we use the **z-score**. The **z-score** (or **standardised score**) indicates the position of a particular score in relation to the mean.  $z$ -scores are a very important statistical measure and later in the chapter some of their uses will be explained.

A  $z$ -score of 0 indicates that the score obtained is equal to the mean, a negative  $z$ -score indicates that the score is below the mean and a positive  $z$ -score indicates a score above the mean.

The  $z$ -score measures the distance from the mean in terms of the standard deviation. A score that is exactly one standard deviation above the mean has a  $z$ -score of 1. A score that is exactly one standard deviation below the mean has a  $z$ -score of  $-1$ .

To calculate a  $z$ -score we use the formula:

$$z = \frac{x - \bar{x}}{s}$$

where  $x$  is the score,  $\bar{x}$  is the mean and  $s$  is the standard deviation.

### WORKED EXAMPLE 1

**In an IQ test the mean IQ is 100 and the standard deviation is 15. Dale's test results give an IQ of 130. Calculate this as a  $z$ -score.**

#### THINK

- 1 Write the formula.
- 2 Substitute for  $x$ ,  $\bar{x}$  and  $s$ .
- 3 Calculate the  $z$ -score.

#### WRITE

$$\begin{aligned} z &= \frac{x - \bar{x}}{s} \\ &= \frac{130 - 100}{15} \\ &= 2 \end{aligned}$$

Dale's  $z$ -score is 2, meaning that his IQ is exactly two standard deviations above the mean.

Not all  $z$ -scores will be whole numbers; in fact most will not be whole numbers. A whole number indicates only that the score is an exact number of standard deviations above or below the mean.

## WORKED EXAMPLE 2

eBookplus

A sample of professional basketball players gives the mean height as 192 cm with a standard deviation of 12 cm. Dieter is 183 cm tall. Calculate Dieter's height as a z-score.

Tutorial  
int-2438

Worked example 2

### THINK

- 1 Write the formula.
- 2 Substitute for  $x$ ,  $\bar{x}$  and  $s$ .
- 3 Calculate the z-score.

### WRITE

$$\begin{aligned} z &= \frac{x - \bar{x}}{s} \\ &= \frac{183 - 192}{12} \\ &= -0.75 \end{aligned}$$

The negative z-score in worked example 2 indicates that Dieter's height is below the mean but, in this case, by less than one standard deviation.

When examining z-scores, care must be taken to use the appropriate value for the standard deviation. If examining a population, the population standard deviation ( $\sigma_n$ ) should be used and if a sample has been taken, the sample standard deviation ( $\sigma_{n-1}$  or  $s_n$ ) should be used.

*Remember:* Your graphics calculator displays all of this information once data is stored and calculated using the statistics function.

## WORKED EXAMPLE 3

To obtain the average number of hours study done by students in her class per week, Kate surveys 20 students and obtains the following results.

12 18 15 14 9 10 13 12 18 25  
15 10 3 21 11 12 14 16 17 20

- a Calculate the mean and standard deviation (correct to 3 decimal places).
- b Robert does 16 hours of study each week. Express this as a z-score based on the above results. (Give your answer correct to 3 decimal places.)

### THINK

#### Method 1: Technology-free

- a
  - 1 Enter the data into your calculator.
  - 2 Obtain the mean from your calculator.
  - 3 Obtain the standard deviation from your calculator using the sample standard deviation.
- b
  - 1 Write the formula.
  - 2 Substitute for  $x$ ,  $\bar{x}$  and  $s$ .
  - 3 Calculate the z-score.

### WRITE

$$\begin{aligned} \bar{x} &= 14.25 \\ s_n &= 4.753 \end{aligned}$$

$$\begin{aligned} z &= \frac{x - \bar{x}}{s} \\ &= \frac{16 - 14.25}{4.753} \\ &= 0.368 \end{aligned}$$

#### Method 2: Technology-enabled

- a
  - 1 From the MENU select STAT.



- 2 Delete any existing data and enter the scores from worked example 3 in **List 1**.



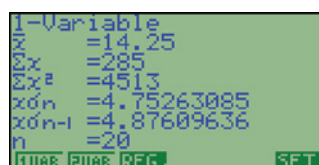
- 3 Press **F2** (CALC). You may need to press **F6** first for more options.



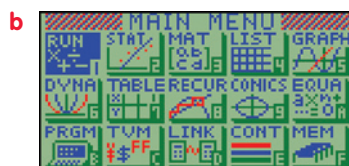
- 4 Press **F6** (SET). Check that **1Var Xlist** is set to **List 1** and **1Var Freq** is set to **1**.



- 5 Press **EXE** to return to the previous screen, and then press **F1** (1VAR) to display the summary statistics.



- b** 1 Press **MENU** and then select **RUN**.



- 2 Press **OPTN** **F6** for more options and then **F3** (PROB).



- 3 Again press **F6** for more options and then **F4** **t()**. This is the z-score function, so enter 16, close brackets and press **EXE**.



### REMEMBER

1. A data set is normally distributed if it is symmetrical about the mean.
2. The graph of a normally distributed data set is a bell-shaped curve that is symmetrical about the mean. In such a distribution the mean, median and mode are equal.
3. A z-score is used to measure the position of a score in a data set relative to the mean.
4. The formula used to calculate a z-score is:

$$z = \frac{x - \bar{x}}{s}$$

where  $x$  is the score,  $\bar{x}$  is the mean, and  $s$  is the standard deviation.



## eBookplus

Digital doc  
SKILLSHEET 11.1  
doc-1408

Finding the  
mean

## eBookplus

Digital doc  
SKILLSHEET 11.2  
doc-1409

Finding the  
standard  
deviation

## eBookplus

Digital doc  
SKILLSHEET 11.3  
doc-1410

Choosing the  
appropriate  
standard  
deviation

1 **WE1** In a Maths exam the mean score is 60 and the standard deviation is 12. Chifune's mark is 96. Calculate her mark as a z-score.

2 In an English test the mean score was 55 with a standard deviation of 5. Adrian scored 45 on the English test. Calculate Adrian's mark on the test as a z-score.

3 **a** In a normal distribution the mean is 32 and the standard deviation 6. Convert a score of 44 to a z-score.

**b** In a normal distribution the mean is 1.2 and the standard deviation is 0.3. Convert a score of 0.6 to a z-score.

**c** The mean of a distribution is 254 and the standard deviation is 39. Write a score of 214 as a standardised score, correct to 2 decimal places.

**d** The mean mark on an exam is 62 and the standard deviation is 9.5. Convert a mark of 90 to a z-score. (Give your answer correct to 2 decimal places.)

4 Tracy is a nurse and samples the mass of 50 newborn babies born in the hospital in which she works. She finds that the mean mass is 3.5 kg, with a standard deviation of 0.4 kg. What would be the standardised score of a baby whose birth mass was:

- a** 3.5 kg?      **b** 3.9 kg?  
**c** 2.7 kg?      **d** 4.7 kg?  
**e** 3.1 kg?



5 Ricky finds that the mean number of hours spent watching television each week by Year 12 students is 10.5 hours, with a standard deviation of 3.2 hours. How many hours of television is watched by a person who has a standardised score of:

- a** 0?      **b** 1?      **c** 2?      **d** -1?      **e** -3?

6 **WE2** IQ tests have a mean of 100 and a standard deviation of 15. Calculate the z-score for a person with an IQ of 96. (Give your answer correct to 2 decimal places.)

7 The mean time taken for a racehorse to run 1 km is 57.69 s, with a standard deviation of 0.36 s. Calculate the z-score of a racehorse that runs 1 km in 58.23 s.

8 In a major exam every subject has a mean score of 60 and a standard deviation of 12.5. Clarissa obtains the following marks on her exams. Express each as a z-score.

- a** English 54      **b** Maths 78      **c** Biology 61      **d** Geography 32      **e** Art 95

9 The mean time for athletes over 100 m is 10.3 s, with a standard deviation of 0.14 s. What time would correspond to a z-score of:

- a** 0?      **b** 2?      **c** 0.5?      **d** -3?      **e** -0.35?      **f** 1.6?

10 **a** Explain what is meant by a z-score of 1.      **b** Explain what is meant by a z-score of -2.

11 **WE3** The length of bolts being produced by a machine needs to be measured. To do this, a sample of 20 bolts are taken and measured. The results (in mm) are given below.

20 19 18 21 20 17 19 21 22 21  
17 17 21 20 17 19 18 22 22 20

**a** Calculate the mean and standard deviation of the distribution.

**b** A bolt produced by the machine is 22.5 mm long. Express this result as a z-score. (Give your answer correct to 2 decimal places.)

- | Amount (\$) | Class centre | Frequency |
|-------------|--------------|-----------|
| 0—< 20      |              | 2         |
| 20—< 40     |              | 8         |
| 40—< 60     |              | 19        |
| 60—< 80     |              | 15        |
| 80—< 100    |              | 6         |

- 13 a** In a distribution, the mean is 50 and the standard deviation is 10. What score corresponds to a z-score of 0?
- b** In a distribution the mean score is 60. If a mark of 76 corresponds to a standardised score of 2, what is the standard deviation?

- Digital doc**  
WorkSHEET 11.1  
doc-1411

Subject	Luke's mark	Mean	Standard deviation
English	72	60	12
Maths	72	55	13
Biology	76	64	8
Computing studies	60	70	5
Visual arts	60	50	15
Music	50	58	10

Chapter 11 The normal distribution | 347

### Further development

- 19** To qualify for an elite sports camp an applicant must be able to run 3 km with a  $z$ -score less than  $-1.5$  in comparison with the general population.  
For males the mean time is 15 min with a standard deviation of 1 min 12 sec  
For females the mean time is 16 min 30 sec with a standard deviation of 1 min 30 sec.  
Calculate the qualification time for both males and females.
- 20** Explain why each of the following statements is true.
- a** In any normally distributed data set the  $z$ -score of the mean is equal to 0.
  - b** In any normally distributed data set a person who performs one standard deviation better than the mean will have a  $z$ -score of 1.
- 21** The following set of scores represent the ages of the members of a choir.  
23 45 32 70 61 44 49 22 36 20
- a** Find the mean and standard deviation of the data set.
  - b** Convert each of the scores to a  $z$ -score.
  - c** What is the mean and the standard deviation of the  $z$ -scores found in part (b)?
  - d** Explain why you can give the answer to part (c) without doing any calculations.
- 22** Ten people are surveyed for their weekly income and the results are below.  
\$485, \$623, \$312, \$1754, \$86, \$740, \$595, \$700, \$490, \$642
- a** Find the mean and standard deviation of the data set.
  - b** Which scores in the data set could be identified as outliers?
  - c** Find the  $z$ -score for any scores which you have identified as outliers.
- 23** John obtained a mark of 76 on an exam. The standard deviation on this test was 8 and John's score corresponded to a  $z$ -score of 1.  
Find the  $z$ -score for each of the following students.
- a** Anthea, who scored 60
  - b** Bianca, who scored 88
  - c** Carl, who scored 68
  - d** Demetria, who scored 40
- 24** The mean height for an adult male is 174 cm with a standard deviation of 8.4 cm. Barney, who is 165 cm tall, is generally considered by most people to be short. Fred, who is 187 cm tall, is generally considered by most people to be tall. Compare Fred's tallness with Barney's shortness.

## 11B Comparison of scores

An important use of  $z$ -scores is to compare scores from different data sets. Suppose that in your Maths exam your result was 74 and in English your result was 63. In which subject did you achieve the better result?

It may appear, at first glance, that the Maths result is better, but this does not take into account the difficulty of the test. A mark of 63 on a difficult English test may in fact be a better result than 74 if it was an easy Maths test.

The only way that we can fairly compare the results is by comparing each result with its mean and standard deviation. This is done by converting each result to a  $z$ -score.

$$\begin{aligned}\text{If for Maths } \bar{x} = 60 \text{ and } s = 12, \text{ then} \quad z &= \frac{x - \bar{x}}{s} \\ &= \frac{74 - 60}{12} \\ &= 1.17\end{aligned}$$
$$\begin{aligned}\text{And if for English } \bar{x} = 50 \text{ and } s = 8, \text{ then } z &= \frac{x - \bar{x}}{s} \\ &= \frac{63 - 50}{8} \\ &= 1.625\end{aligned}$$



The English result is better because the higher  $z$ -score shows that the 63 is higher in comparison to the mean of each subject.

#### WORKED EXAMPLE 4

**eBookplus**

**Tutorial**  
int-2439

Worked example 4

Janine scored 82 in her Physics exam and 78 in her Chemistry exam.

In Physics,  $\bar{x} = 62$  and  $s = 10$ , while in Chemistry,  $\bar{x} = 66$  and  $\sigma_n = 5$ .

- a Write both results as a standardised score.
- b Which is the better result? Explain your answer.

#### THINK

- 1 Write the formula for each subject.
  - 2 Substitute for  $x$ ,  $\bar{x}$  and  $s$ .
  - 3 Calculate each  $z$ -score.
- b Explain that the subject with the highest  $z$ -score is the better result.

#### WRITE

- a Physics:  $z = \frac{x - \bar{x}}{s}$       Chemistry:  $z = \frac{x - \bar{x}}{s}$   
$$= \frac{82 - 62}{10}$$
$$= 2$$
  
Chemistry:  $z = \frac{78 - 66}{5}$ 
$$= 2.4$$
- b The Chemistry result is better because of the higher  $z$ -score.

In each example the circumstances must be read carefully to see whether a higher or lower  $z$ -score is better. For example, if we were comparing times for runners over different distances, the lower  $z$ -score would be the better one.

#### WORKED EXAMPLE 5

In international swimming the mean time for the men's 100 m freestyle is 50.46 s with a standard deviation of 0.6 s. For the 200 m freestyle, the mean time is 1 min 51.4 s with a standard deviation of 1.4 s. Sam's best time is 49.92 s for 100 m and 1 min 49.3 for 200 m. At a competition Sam can enter only one of these events. Which event should he enter?

#### THINK

- 1 Write the formula for both events.
- 2 Substitute for  $x$ ,  $\bar{x}$  and  $s$ . (For 200 m convert time to seconds.)
- 3 Calculate the  $z$ -scores.
- 4 The best event is the one with the lower  $z$ -score.

#### WRITE

$$\begin{aligned} 100 \text{ m: } z &= \frac{x - \bar{x}}{s} & 200 \text{ m: } z &= \frac{x - \bar{x}}{s} \\ &= \frac{49.92 - 50.46}{0.6} & &= \frac{109.3 - 111.4}{1.4} \\ &= -0.9 & &= -1.5 \end{aligned}$$

The  $z$ -score for 200 m is lower, indicating that Sam's time is further below the mean and that this is the event that he should enter.

#### REMEMBER

1. Scores can be compared by their  $z$ -scores as they compare the score with the mean and the standard deviation.
2. Read each question carefully to see if a higher or lower  $z$ -score is a better outcome.

## Comparison of scores

eBookplus

Digital doc

EXCEL Spreadsheet

doc-1412

One variable

statistics

- 1 **WE4** Ken's English mark was 75 and his Maths mark was 72. In English the mean was 65 with a standard deviation of 8, while in Maths the mean mark was 56 with a standard deviation of 12.
- Convert the mark in each subject to a  $z$ -score.
  - In which subject did Ken perform better? Explain your answer.
- 2 In the first Maths test of the year the mean mark was 60 and the standard deviation was 12. In the second test the mean was 55 and the standard deviation was 15. Barbara scored 54 in the first test and 50 in the second test. In which test did Barbara do better? Explain your answer.
- 3 **MC** The table below shows the mean and standard deviation in four subjects.

Subject	Mean	Standard deviation
English	60	12
Maths	65	8
Biology	62	16
Geography	52	7.5

Kelly's marks were English 66, Maths 70, Biology 50 and Geography 55. In which subject did Kelly achieve her best result?

- A** English      **B** Maths      **C** Biology      **D** Geography
- 4 **MC** The table below shows the mean and standard deviation of house prices in four Australian cities. The table also shows the cost of building the same three-bedroom house in each of the cities.

City	Mean	Standard deviation	Cost
Sydney	\$230 000	\$30 000	\$215 000
Melbourne	\$215 000	\$28 000	\$201 000
Adelaide	\$185 000	\$25 000	\$160 000
Brisbane	\$190 000	\$20 000	\$165 000

In which city is the standardised cost of building the house least?

- A** Sydney      **B** Melbourne  
**C** Adelaide      **D** Brisbane
- 5 **WES** Karrie is a golfer who scored 70 on course A, which has a mean of 72 and a standard deviation of 2.5. On course B, Karrie scores 69. The mean score on course B is 72 and the standard deviation is 4. On which course did Karrie play the better round? (In golf the lower score is better.)
- 6 Steve is a marathon runner. On the Olympic course in Sydney the mean time is 2 hours and 15 minutes with a standard deviation of 4.5 minutes. On Athens' Olympic course the mean time is 2 hours and 16 minutes with a standard deviation of 3 minutes. In Sydney Steve's time was 2 hours 17 minutes and in Athens his time was 2 hours 19 minutes.
- Write both times as a  $z$ -score.
  - Which was the better performance? Explain your answer.



- 7 **MC** The table below shows the mean and standard deviation of times in the 100 m by the same group of athletes on four different days. It also shows Matt's time on each of these days.

Day	Mean	Standard deviation	Matt's time
8 Jan.	10.21	0.15	10.12
15 Jan.	10.48	0.28	10.30
22 Jan.	10.14	0.09	10.05
29 Jan.	10.22	0.12	10.11

On what day did Matt give his best performance?

- A 8 Jan.      B 15 Jan.      C 22 Jan.      D 29 Jan.

- 8 **MC** In which of the following subjects did Alyssa achieve her best standardised result?

Subject	Alyssa's mark	Mean	Standard deviation
English	54	60	12
Maths	50	55	15
Biology	60	65	8
Music	53	62	9

- A English      B Maths      C Biology      D Music

- 9 Shun Mei received a mark of 64 on her Maths exam and 63 on her Chemistry exam. To determine how well she actually did on the exams, Shun Mei sampled 10 people who sat for the same exams and the results are shown below.

Maths: 56 45 82 90 41 32 65 60 55 69

Chemistry: 55 63 39 92 84 46 47 50 58 62

- a Calculate the mean and standard deviation for Shun Mei's sample in each subject.  
b By converting each of Shun Mei's marks to  $z$ -scores, state the subject in which she performed best.

- 10 Ricardo scored 85 on an entrance test for a job. The test has a mean score of 78 and a standard deviation of 8. Kory sits a similar exam and scores 27. In this exam the mean is 18 and the standard deviation is 6. Who is better suited for the job? Explain your answer.

- 11 a Cynthia scored a mark of 65 in English where the mean was 55 and the standard deviation is 8. In Maths Cynthia scored 66 where the mean was 52 and the standard deviation 10. Convert the mark in each subject to a  $z$ -score.  
b In which subject did Cynthia achieve her best result?

### Further development

- 12 Eamon is a swimmer who can swim 100 m in 55 s at the state championships. The mean time for his age group at the championships is 58 s with a standard deviation of 1.5 s. Stephanie swims 100 m in 61.5 s, when the mean is 63 s with a standard deviation of 1 s. Whose performance is better? Explain your answer.
- 13 The mean price of a litre of milk is \$1.30 with a standard deviation of 10c, and the mean price of a loaf of bread is \$3.10 with a standard deviation of 25c. Joe's convenience store charges \$1.55 for a litre of milk and \$3.70 for a loaf of bread. Determine which of the milk and the bread is relatively more expensive.
- 14 The average number of tourists visiting Cairns each year is 240 000 with a standard deviation of 15 000. During 2009 the number of visitors fell to 205 000.
- a Find this number of visitors as a standardised score. (Give your answer correct to 2 decimal places.)  
b What does this tell you about the number of visitors for 2009?  
c Suggest some possible reasons for this result.

- 15** The mean number of visitors on a weekday to a theme park is 2000 and on the weekend is 3500. The number of visitors on a particular Saturday is 4550 and this corresponds to a  $z$ -score of 1.6.
- Find the standard deviation for the number of weekend visitors.
  - On weekdays the standard deviation is 345. Calculate the equivalent number of weekday visitors.
- 16** A liquid dispenser on a production line is responsible for pouring motor oil into 1.5-L containers. The mean amount poured is 1.5 L and the standard deviation is 8 mL. To test that the each dispenser is working properly, a sample of five containers is taken and the contents of each measured. If any container has a  $z$ -score of greater than 2.5 or less than  $-1.5$ , the machine is closed down until further testing takes place. From a sample of five the following results were obtained: 1490 mL, 1508 mL, 1519 mL, 1486 mL and 1500 mL. Determine if the machine needs to be closed down for further testing.
- 17** Jerry and Sammy are brothers. Jerry lives in London where the mean January temperature is  $8^{\circ}\text{C}$  and the standard deviation is  $1.5^{\circ}\text{C}$ . Sammy lives in Sydney where the mean January temperature is  $27^{\circ}\text{C}$  with a standard deviation of  $2.8^{\circ}\text{C}$ . During a telephone call Jerry comments on how warm it is in London when the temperature is  $11^{\circ}\text{C}$  and Sammy says that it is cool in Sydney at  $22^{\circ}\text{C}$ . Explain what each of the two brothers mean by their comments, and compare the warmth in London with Sydney's cool weather.

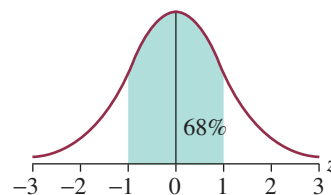
**eBookplus**

**Digital doc**  
Investigation  
**Comparison**  
of subjects  
doc-1413

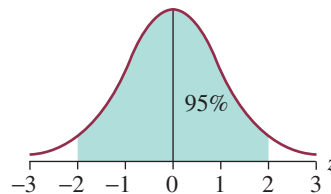
## 11C Distribution of scores

In any normal distribution, the percentage of scores that lie within a certain number of standard deviations of the mean is always the same, provided that the sample is large enough. This is true irrespective of the values of the mean and standard deviation.

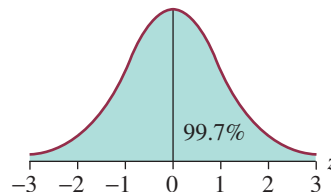
In any normal distribution, approximately 68% of the values will lie within one standard deviation of the mean. This means 68% of scores will have a  $z$ -score between  $-1$  and  $1$ . This can be shown on a normal curve as:



Approximately 95% of the values lie within 2 standard deviations, or have a  $z$ -score of between  $-2$  and  $2$ .



Approximately 99.7% of scores lie within 3 standard deviations, or have a  $z$ -score that lies between  $-3$  and  $3$ .



If we know that a random variable is approximately normally distributed, and we know its mean and standard deviation, then we can use this rule to quickly make some important statements about the way in which the data values are distributed.

### WORKED EXAMPLE 6

eBookplus

Tutorial  
int-2440

Worked example 6

Experience has shown that the scores obtained on a commonly used IQ test can be assumed to be normally distributed with a mean of 100 and a standard deviation of 15. Approximately what percentage of the distribution lies:

- a** between 85 and 115? **b** between 70 and 130? **c** between 55 and 145?

#### THINK

- a** 1 Calculate the  $z$ -scores for 85 and 115.
- 2 68% of scores have a  $z$ -score between  $-1$  and  $1$ .
- b** 1 Calculate the  $z$ -scores for 70 and 130.
- 2 95% of scores have a  $z$ -score between  $-2$  and  $2$ .
- c** 1 Calculate the  $z$ -scores for 55 and 145.
- 2 99.7% of scores have a  $z$ -score between  $-3$  and  $3$ .

#### WRITE

$$\mathbf{a} \quad z = \frac{85-100}{15} = -1 \qquad z = \frac{115-100}{15} = 1$$

68% of the scores will lie between 85 and 115.

$$\mathbf{b} \quad z = \frac{70-100}{15} = -2 \qquad z = \frac{130-100}{15} = 2$$

95% of the scores will lie between 70 and 130.

$$\mathbf{c} \quad z = \frac{55-100}{15} = -3 \qquad z = \frac{145-100}{15} = 3$$

99.7% will lie between 55 and 145.

We can also make statements about the percentage of scores that lie in the tails of the distribution by using the symmetry of the distribution and remembering that 50% of scores will have a  $z$ -score greater than 0 and 50% will have a  $z$ -score less than 0.

### WORKED EXAMPLE 7

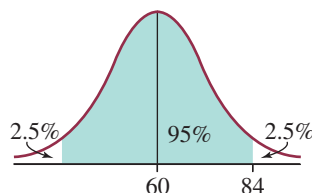
In an exam  $\bar{x} = 60$  and  $s = 12$ . What percentage of candidates in the exam scored above 84?

#### THINK

- 1 Calculate 84 as a  $z$ -score.
- 2 Draw a sketch showing 95% of  $z$ -scores lie between  $-2$  and  $2$ .
- 3 5% of  $z$ -scores therefore lie outside this range. Half of these scores lie below  $-2$  and half are above  $2$ .
- 4 Give a written answer.

#### WRITE

$$z = \frac{x - \bar{x}}{s} = \frac{84 - 60}{12} = 2$$



2.5% of scores are greater than 84.

Some important terminology is used in connection with this rule. We can say that if 95% of scores have a  $z$ -score between  $-2$  and  $2$ , then if one member of the population is chosen, that member will *very probably* have a  $z$ -score between  $-2$  and  $2$ .



If 99.7% of the population has a  $z$ -score between  $-3$  and  $3$ , then if one member of that population is chosen, that member will *almost certainly* have a  $z$ -score between  $-3$  and  $3$ .

### WORKED EXAMPLE 8

A machine produces tyres that have a mean thickness of 12 mm, with a standard deviation of 1 mm. If one tyre that has been produced is chosen at random, within what limits will the thickness of the tyre:

- a very probably lie?
- b almost certainly lie?

#### THINK

- a
  - 1 Tyre thickness will *very probably* have a  $z$ -score between  $-2$  and  $2$ .
  - 2 A  $z$ -score of  $-2$  corresponds to a tyre of 10 mm thickness.
  - 3 A  $z$ -score of  $2$  corresponds to a tyre of 14 mm thickness.
- b
  - 1 Tyre thickness will *almost certainly* have a  $z$ -score between  $-3$  and  $3$ .
  - 2 A  $z$ -score of  $-3$  corresponds to a tyre of 9 mm thickness.
  - 3 A  $z$ -score of  $3$  corresponds to a tyre of 15 mm thickness.

#### WRITE

<p>a If <math>z = -2</math></p> $x = \bar{x} - 2s$ $= 12 - 2 \times 1$ $= 10$	<p>If <math>z = 2</math></p> $x = \bar{x} + 2s$ $= 12 + 2 \times 1$ $= 14$
---	--

A tyre chosen will *very probably* have a thickness of between 10 and 14 mm.

<p>b If <math>z = -3</math></p> $x = \bar{x} - 3s$ $= 12 - 3 \times 1$ $= 9$	<p>If <math>z = 3</math></p> $x = \bar{x} + 3s$ $= 12 + 3 \times 1$ $= 15$
--	--

A tyre chosen will *almost certainly* have a thickness of between 9 and 15 mm.

Because it is almost certain that a member of the data set will lie within three standard deviations of the mean, if a possible member of the data set is found to be outside this range one should suspect a problem.

For example, if a machine is set to deposit 200 mL of liquid into a bottle, with a standard deviation of 5 mL, and then a bottle is found to have contents of 220 mL, one would expect there to be a problem with the settings on the machine.

This knowledge of  $z$ -scores is then used in industry by the quality control department. In the above example a sample of bottles would be tested and the  $z$ -scores recorded. The percentage of  $z$ -scores between  $-1$  and  $1$ ,  $-2$  and  $2$ , and  $-3$  and  $3$  are checked against the above rule. If these percentages are not correct, the machinery needs to be checked for faults.

### REMEMBER

1. In a normal distribution:
  - 68% of scores will have a  $z$ -score between  $-1$  and  $1$
  - 95% of scores will have a  $z$ -score between  $-2$  and  $2$
  - 99.7% of scores will have a  $z$ -score between  $-3$  and  $3$ .
2. The symmetry of the normal distribution allows us to make calculations about the percentage of scores lying within certain limits.
3. If a member of a normally distributed population is chosen, it will:
  - *very probably* have a  $z$ -score between  $-2$  and  $2$
  - *almost certainly* have a  $z$ -score between  $-3$  and  $3$ .
4. Any score further than three standard deviations from the mean indicates that there may be a problem with the data set.

## Distribution of scores

- 1 **WE6** The temperature on a January day in a city is normally distributed with a mean of  $26^{\circ}$  and a standard deviation of  $3^{\circ}$ . What percentage of January days lie between:
  - a  $23^{\circ}$  and  $29^{\circ}$ ?
  - b  $20^{\circ}$  and  $32^{\circ}$ ?
  - c  $17^{\circ}$  and  $35^{\circ}$ ?
- 2 The marks of students sitting for a major exam are normally distributed with mean equal to 57 and standard deviation equal to 13. What percentage of marks on the exam were between:
  - a 44 and 70?
  - b 31 and 83?
  - c 18 and 96?
- 3 The mean thickness of bolts produced by a machine is 2.3 mm, with a standard deviation of 0.04 mm. What percentage of bolts will have a thickness between 2.22 mm and 2.38 mm?
- 4 **WE7** Experience has shown that the scores obtained on a commonly used IQ test can be assumed to be normally distributed with a mean of 100 and a standard deviation of 15. What percentage of scores lie above 115?
- 5 The heights of young women are normally distributed with a mean  $\bar{x} = 160$  cm and a standard deviation equal to 8 cm. What percentage of the women would you expect to have heights:
  - a between 152 and 168 cm?
  - b greater than 168 cm?
  - c less than 136 cm?
- 6 The age at which women give birth to their first child is normally distributed with  $\bar{x} = 27.5$  years and  $\sigma_n = 3.2$  years. From these data we can conclude that about 95% of women have their first child between what ages?
- 7 Fill in the blanks in the following statements. For any normal distribution:
  - a 68% of the values have a  $z$ -score between \_\_\_\_ and \_\_\_\_
  - b \_\_\_\_% of the values have a  $z$ -score between  $-2$  and  $2$
  - c \_\_\_\_% of the values have a  $z$ -score between \_\_\_\_ and \_\_\_\_.
- 8 **MC** Medical tests indicate that the amount of an antibiotic needed to destroy a bacterial infection in a patient is normally distributed with  $\bar{x} = 120$  mg and  $\sigma_n = 15$  mg. The percentage of patients who would require more than 150 mg to clear the infection is:
  - A 0.15%
  - B 2.5%
  - C 5%
  - D 95%
- 9 **MC** The mean mark on a test is 55, with a standard deviation of 10. The percentage of students who achieved a mark between 65 and 75 is:
  - A 13.5%
  - B 22.5%
  - C 34%
  - D 95%
- 10 In a factory, soft drink is poured into cans such that the mean amount of soft drink is 500 mL with a standard deviation of 2 mL. Cans with less than 494 mL of soft drink are rejected and not sold to the public. What percentage of cans are rejected?
- 11 The distribution of IQ scores for the inmates of a certain prison is approximately normal with a mean of 85 and a standard deviation of 15.
  - a What percentage of this prison population have an IQ of 100 or higher?
  - b If someone with an IQ of 70 or less can be classified as mentally disabled, what percentage of the prison population could be classified as mentally disabled?
- 12 The distribution of blood pressures (systolic) among women of similar ages is normal with a mean of 120 (mm of mercury) and a standard deviation of 10 (mm of mercury). Determine the percentage of women with a systolic blood pressure:
  - a between 100 and 140
  - b greater than 130
  - c between 120 and 130
  - d between 90 and 110
  - e between 110 and 150.
- 13 **WE8** The mass of packets of chips is normally distributed with  $\bar{x} = 100$  g and  $\sigma_n = 2.5$  g. If I purchase a packet of these chips, between what limits will the mass of the packet:
  - a very probably lie?
  - b almost certainly lie?

- 14** The heights of army recruits are normally distributed about a mean of 172 cm and a standard deviation of 4.5 cm. A volunteer is chosen from the recruits. The height of the volunteer will very probably lie between what limits?
- 15** A machine is set to deposit a mean of 500 g of washing powder into boxes with a standard deviation of 10 g. When a box is checked, it is found to have a mass of 550 g. What conclusion can be drawn from this?
- 16** The average mass of babies is normally distributed with a mean of 3.8 kg and a standard deviation of 0.4 kg. A newborn baby will almost certainly have a mass between what limits?

### Further development

- 17** The force required to tear a piece of fabric is normally distributed. A sample of the fabric is tested and it is found that the mean breaking force was 6 N with a standard deviation of 0.15 N.  
Write down the values between which:
- a** 68% of the scores would lie
  - b** 95% of the scores would lie
  - c** 99.7% of the scores would lie.
- 18** **MC** A distribution of scores is normally distributed with a mean of 42. It is known that 68% of scores lie between 32 and 52.  
The standard deviation of the data set is:
- |             |             |
|-------------|-------------|
| <b>A</b> 5  | <b>B</b> 10 |
| <b>C</b> 15 | <b>D</b> 20 |
- 19** The salaries of a group of business people is normally distributed. The mean salary is \$65 000 with a standard deviation of \$5000.  
Find the percentage of business people who:
- a** have a salary less than \$70 000
  - b** have a salary less than \$50 000
  - c** have a salary more than \$55 000
  - d** have a salary between \$50 000 and \$75 000.
- 20** In a school the number of days absent by students is normally distributed. The mean number of days absent is 15 days with a standard deviation of 4 days.  
Find the percentage of students who have:
- a** more than 19 days off per year
  - b** fewer than 3 days off per year
  - c** more than 7 days off per year
  - d** between 11 and 23 days off per year.
- 21** The volume of soft drink poured into bottles is measured by the quality control department, and the results are found to be normally distributed. The department found that the mean was 375 mL and the standard deviation was 5 mL.  
From a sample of 400 bottles, how many would be expected to have a volume of:
- a** more than 375 mL
  - b** more than 370 mL
  - c** more than 380 mL
  - d** less than 365 mL?
- 22** The masses of each ball bearing in a batch are normally distributed. The mean mass of a ball bearing is 25 g and the standard deviation is 1 g.  
The manufacturer rejects all ball bearings that are less than 23 g in mass and more than 26 g in mass. From a sample of 2500 ball bearings, how many ball bearings will be rejected?

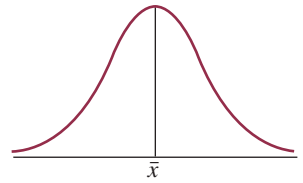
# SUMMARY

## z-scores

- A data set is normally distributed if it is symmetrical about the mean.
- A  $z$ -score measures the position of a score relative to the mean and standard deviation.
- A  $z$ -score is found using the formula

$$z = \frac{x - \bar{x}}{s}$$

where  $x$  is the score,  $\bar{x}$  is the mean, and  $s$  is the standard deviation.



## Comparison of scores

- Standardising both scores best compares scores from different data sets.
- When comparing exam marks, the highest  $z$ -score is the best result.

## Distribution of scores

- A data set that is normally distributed will be symmetrical about the mean.
- 68% of scores will have a  $z$ -score of between  $-1$  and  $1$ .
- 95% of scores will have a  $z$ -score between  $-2$  and  $2$ . A score chosen from this data set will very probably lie in this range.
- 99.7% of scores will have a  $z$ -score of between  $-3$  and  $3$ . A score chosen from the data set will almost certainly lie within this range.

# CHAPTER REVIEW

## MULTIPLE CHOICE

- 1 The mean time for 12-year-old boys to swim 50 m is 50.5 s with a standard deviation of 4.2 s. Kyle swims 50 m in 44.2 s. Kyle's time as a standardised score is:
- A -6.3
  - B -1.5
  - C 1.5
  - D 6.3



- 2 A teacher converts the marks on every test that she gives her class to a standardised score. On a test the mean mark was 50 and the standard deviation was 10. Adam's standardised score on the test was  $-0.6$ . Adam's mark on the test was:
- A 40
  - B 44
  - C 56
  - D 60

- 3 The details of Andrea's half-yearly exams are shown in the table below.

Subject	Andrea's mark	Mean	Standard deviation
English	65	50	12
Maths	62	52	6
History	75	58	15
Geography	50	44	4

In which subject did Andrea achieve her best result?

- A English
- B Maths
- D Geography
- C History

- 4 The details of Brett's half-yearly exams are shown in the table below.

Subject	Brett's mark	Mean	Standard deviation
English	40	50	12
Maths	48	52	6
History	49	58	15
Geography	42	44	4

In which subject did Brett achieve his best result?

- A English
  - B Maths
  - C History
  - D Geography
- 5 A data set is normally distributed with  $\bar{x} = 25$  and  $s = 2.5$ . The percentage of scores that will lie in the range 25 to 30 is:
- A 34%
  - B 47.5%
  - C 68%
  - D 95%
- 6 A fishing boat catches a load of fish and finds the mass of each fish. The masses of the fish are normally distributed with a mean of 800 g and a standard deviation of 75 g. If a fish is chosen from the catch, its mass will almost certainly lie between:
- A 725 g and 875 g
  - B 650 g and 950 g
  - C 575 g and 1025 g
  - D 800 g and 1025 g





## SHORT ANSWER

- 1** Measurements of the amount of acid in a certain chemical are made. The results are normally distributed such that the mean is 6.25% and the standard deviation is 0.25%. Harlan gets a reading of 5.75%. What is Harlan's reading as a  $z$ -score?



- 2** A set of scores is normally distributed such that  $\bar{x} = 15.3$  and  $s = 5.2$ . Convert each of the following members of the distribution to  $z$ -scores.

- a** 15.3
- b** 20.5
- c** 4.9
- d** 30.9
- e** 10.1

- 3** On an exam the results are normally distributed with a mean of 58 and a standard deviation of 7.5. Jennifer scored a mark of 72 on the exam. Convert Jennifer's mark to a  $z$ -score, giving your answer correct to 2 decimal places.

- 4** A set of scores is normally distributed with a mean of 2.8 and a standard deviation of 0.6. Convert each of the following members of the data set to  $z$ -scores, correct to 2 decimal places.

- a** 2.9
- b** 3.9
- c** 1
- d** 1.75
- e** 1.6

- 5** Anji conducts a survey on the water temperature at her local beach each day for a month. The results (in  $^{\circ}\text{C}$ ) are shown below.

20 21 19 22 21 18 17 23 17 16  
 20 21 18 22 17 16 20 20 22 19  
 22 20 20 20 21  
 21 22 23 24 20

- a** Find the mean and standard deviation of the scores.

- b** Find the highest and lowest temperatures in the data set and express each as a  $z$ -score.

- 6** The table below shows the length of time for which a sample of 100 light bulbs will burn.



Length of time (hours)	Class centre	Frequency
0–< 500		3
500–< 1000		28
1000–< 1500		59
1500–< 2000		10

- a** Find the mean and standard deviation for the data set.

- b** A further sample of five light bulbs are chosen. The length of time for which each light bulb burned is given below. Convert each of the following to a standardised score.

- i** 1000 hours
- ii** 1814 hours
- iii** 256 hours
- iv** 751 hours
- v** 2156 hours

- 7** Betty sat exams in both Physics and Chemistry. In Physics the exam results showed a mean of 48 and a standard deviation of 12, while in Chemistry the mean was 62 with a standard deviation of 9.
- a** Betty scored 66 in Physics. Convert this result to a z-score.
  - b** Betty scored 71 in Chemistry. Convert this result to a z-score.
  - c** In which subject did Betty achieve the better result? Explain your answer.



- 8** In Geography Carlos scored a mark of 56, while in Business studies he scored 58. In Geography  $\bar{x} = 64$  and  $s = 10$ . For Business studies  $\bar{x} = 66$  and  $s = 15$ .
- a** Convert each mark to a standardised score.
  - b** In which subject did Carlos achieve the better result?
- 9** A psychologist records the number of errors made on a series of tests. On a literacy test the mean number of errors is 15.2 and the standard deviation is 4.3. On the numeracy test the mean number of errors is 11.7 with a standard deviation of 3.1. Barry does both tests and makes 11 errors on the literacy test and 8 errors on the numeracy test. In which test did Barry do better? Explain your answer.

- 10** A data set is normally distributed with a mean of 40 and a standard deviation of 8. What percentage of scores will lie in the range:
- a** 32 to 48?
  - b** 24 to 56?
  - c** 16 to 64?
- 11** The value of sales made on weekdays at a store appears to be normally distributed with a mean of \$1560 and a standard deviation of \$115. On what percentage of days will the days' sales lie between:
- a** \$1445 and \$1675?
  - b** \$1330 and \$1790?
  - c** \$1215 and \$1905?
- 12** A data set is normally distributed with a mean of 56 and a standard deviation of 8. What percentage of scores will:
- a** lie between 56 and 64?
  - b** lie between 40 and 56?
  - c** be less than 40?
  - d** be greater than 80?
  - e** lie between 40 and 80?
- 13** A machine is set to produce bolts with a mean diameter of 5 mm with a standard deviation of 0.1 mm. A bolt is chosen and it is found to have a diameter of 4.5 mm. What conclusion can be drawn about the settings of the machine?



## EXTENDED RESPONSE

- 1 Theresa attempts to review her exam results in Physics and Chemistry. Theresa samples 10 of her friends and finds the following results.
- Physics: 65 64 67 69 72 50 66 66 63 69  
Chemistry: 72 50 69 55 62 68 51 75 78 44
- Find the mean and standard deviation in each subject.
  - Theresa's marks were 65 in Physics and 67 in Chemistry. Convert each to a standardised score.
  - In which subject did Theresa score her best result? Explain your answer.
  - A student is chosen at random from the Physics class. Between what two marks will this person's result very probably lie?
  - If the marks within the class follow a normal distribution, within what two marks will approximately 99.7% of all Chemistry scores lie?
- 2 A machine is set to cut lengths of metal such that the mean length of metal cut is 12.5 cm with a standard deviation of 0.05 cm.
- A piece of metal is measured to have a length of 12.4 cm. Express this as a standardised score.
  - A second piece of metal is measured and found to have a length of 13 cm. What conclusion can be drawn from this measurement?

**eBookplus**

**Digital doc**  
Test Yourself  
doc-1416  
**Chapter 11**

**Are you ready?****Digital docs** (page 342)

- SkillsSHEET 11.1 (doc-1408): Finding the mean.
- SkillsSHEET 11.2 (doc-1409): Finding the standard deviation.
- SkillsSHEET 11.3 (doc-1410): Choosing the appropriate standard deviation.

**11A z-scores****Tutorial**

- **WE2** int-2438: Perform a calculation of a z-score. (page 344)

**Digital docs**

- SkillsSHEET 11.1 (doc-1408): Finding the mean. (page 346)
- SkillsSHEET 11.2 (doc-1409): Finding the standard deviation. (page 346)
- SkillsSHEET 11.3 (doc-1410): Choosing the appropriate standard deviation. (page 346)
- WorkSHEET 11.1 (doc-1411): Apply your knowledge of z-scores. (page 347)

**11B Comparison of scores****Tutorial**

- **WE4** int-2439: Perform a comparison of scores. (page 349)

**Digital docs**

- Spreadsheet (doc-1412): One variable statistics. (page 350)
- Investigation (doc-1413): Comparison of subjects. (page 352)

**11C Distribution of scores****Tutorial**

- **WE6** int-2440: Estimate distribution scores. (page 353)

**Digital docs**

- WorkSHEET 11.2 (doc-1414): Apply your knowledge of distribution scores. (page 356)
- Investigation (doc-1415): Examining a normal distribution. (page 356)

**Chapter review**

- Test Yourself (doc-1416): Take the end-of-chapter test to test your progress. (page 361)

To access eBookPLUS activities, log on to

[www.jacplus.com.au](http://www.jacplus.com.au)